DOCUMENT RESUME

ED 358 116 TM 019 882

AUTHOR	McMullen, Bonita Gibson
TITLE	Quantitative Analysis of Effects in the Classrooms.
SPONS AGENCY	Department of Education, Washington, DC.
PUB DATE	Apr 93
CONTRACT	R117E10152
NOTE	29p.; Paper presented at the Annual Meeting of the American Educational Research Association (Atlanta, GA, April 12-16, 1993).
PUB TYPE	Reports - Research/Technical (143) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Classroom Observation Techniques; *Educational Assessment; *Elementary School Teachers; Elementary Secondary Education; Interviews; *Mathematics Instruction; Mathematics Teachers; Research Methodology; *Secondary School Teachers; *Teaching Methods; Thinking Skills
IDENTIFIERS	*Alternative Assessment; Performance Based Evaluation

ABSTRACT

The framework and results of a study exploring the effects of alternative assessment on mathematics teaching are reported. The instrument for the observation of alternative assessment in the classroom, based on the processes defined by P. L. Peterson (1988), is described along with a discussion of the methodology and the data implications. Eighteen elementary school and secondary school teachers were observed over the 1991-92 school year, as part of the Innovative Mathematics Assessment and Teachers' Classroom Practice project. In addition, results are presented from the collection of informal data, such as teacher interviews, comments, and written summaries describing assessment strategies that worked or did not work in the classes. Study data, along with data from the teacher observation reports, provide evidence that the use of alternative assessments helps mathematics teachers change their teaching strategies, incorporating the use of an emphasis on meaning and understanding, encouragement of student autonomy and persistence, and the direct teaching of higher order cognitive strategies. Six tables present study findings, and three figures illustrate the discussion. (SLD)

Quantitative Analysis of Effects in the Classrooms

Bonita Gibson McMullen
Texas A&M University

Symposium paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA, April 15, 1993.

This work was funded by the United States Department of Education under grant number R117E10152. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the author and do not necessarily reflect the views of the U. S. Department of Education.

ABSRACT

This paper presents the framework and the results of a study exploring the effects of alternative assessment on mathematics teaching. The instrument for the observation of alternative assessment in the classroom, based on the processes defined by Peterson (1988), is described along with a discussion of the methodology and the data implications. In addition, the paper represents the results from the collection of informal data, such as teacher interviews, comments, and written summaries describing assessment strategies that worked or did not work in the classes. This data, along with the data from the teacher observation reports, provide evidence that the use of alternative assessments help mathematics teachers change their teaching strategies; incorporating the use of an emphasis on meaning and understanding; encouragement of student autonomy and persistence, and the direct teaching of higher order cognitive strategies.

Quantitative Analysis of Effects in the Classrooms

Bonita Gibson McMullen, Texas A&M University

Classroom Observation

Much of the recent work in in-service education, (funded by Eisenhower grants, National Science Foundation and other agencies, states, and local school systems) has been aimed at helping teachers to use problem solving, hands-on activities, and other methods which enhance the meaningfulness and understanding of mathematics for all students. Very little of this work has addressed the issue of assessment. Teachers are often taught new methods and are encouraged to use them in class, but are given no help in how to test their students' knowledge when using the new methods. The experience of the Assessment Performance Unit (APU) in England shows that teacher training in methods of assessment can lead to changes in the way they teach (Foxman & Mitchell, 1983). In the APU work, teachers who were trained as administrators of hands-on tests found the test items themselves to be useful classroom teaching activities. This type of experience can lead to the most significant change of all, that is, making teaching and testing "seamless" classroom experiences.

Recent experience in American mathematics education seems to verify that the material tested is the material taught. The evidence from the National Assessment of Educational Progress (NAEP), for example, has shown that computational skills have been the focus for competency tests which have produced textbooks and instructional emphases aimed at developing these skills in students. Teachers have been legitimately concerned that if they "fight the system" and teach higher order thinking, their students would suffer on the computational oriented tests that they are required to pass. Many educators believe that very little change will occur in mathematics curriculum and teaching without a concurrent change in testing, especially in state and national standardized tests that are used to assess and compare school-by-school achievement.

Mathematics educators have called for new approaches to assessment and have provided some direction for changes (Charles & Silver, 1989; Kulm, 1990; NCTM, 1989; Romberg, 1990). Promising work on developing new approaches to mathematics assessment has begun to appear at the national, state and local level (CCSSO, 1989; NAEP, 1987; Pandey, 1990; Stenmark, 1989; Strong, 1990). However, little research has been done on specific approaches to mathematics assessments and how they work with various student populations.

Significance of the Study

The Innovative Mathematics Assessment and Teachers' Classroom Practice (IMAP) focused on the need to help teachers learn to improve the way they evaluate a student's mathematics learning. The major objective was to develop assessment approaches that more accurately reflect the teacher's own goals, and the goals of the NCTM Curriculum and Evaluation Standards (NCTM, 1989) and other national recommendations. A more indirect, but important objective of IMAP was to affect classroom teaching practice. A Core Curriculum, in it's discussion of the new assessment practices in the Netherlands, points out the multidimensional aspects of student assessment and the fact that it is integral to instruction (NCTM, 1992).

Research Questions

Assessment has been identified as one of the most significant challenges facing mathematics education reform. National groups, such as the Mathematical Sciences Education Board, have convened conferences to address the issue. The most important issue, however, has not received much attention. That is, if teachers are trained to use alternative assessment approaches, how will these assessments affect the way teachers teach? And how much will all of this training affect the attitudes of the teachers and the students?

Review of the Literature

In Everybody Counts (Mathematical Sciences Education Board, 1989, p. 70), a report on the future of mathematics education in the United States, the National Research Council asserts that "we must ensure that tests measure what is of value, not just what is easy to test. If we want students to investigate, explore, and discover, assessment must not measure just 'mimicry' mathematics." Instead we need ways to measure the understanding a student obtains during an investigation, exploration, or discovery lesson.

In Mathematics Assessment, it is noted that "new forms of assessment are not goals in and of themselves. The major rationale for diversifying mathematics assessment is the value that the diversification has as a tool for the improvement of our teaching and the students' mathematics learning" (Stenmark, 1991). If the method of assessment gives the teacher the necessary information to recognize what the student understands, then the teacher has the tools to change the way he/she teaches.

Most mathematics teachers believe that higher order thinking is important. In the second International Mathematics Study (IMS), more than 60 percent of U.S. mathematics teachers listed their highest goal as "developing a systematic approach to solving problems and developing an awareness of the importance of mathematics to everyday life" (Crosswhite et. al., 1986). Student performance on the IMS and the recent National Assessment of Educational Progress (NAEP) tests indicate, however, that the aspirations of teachers and the performance of their students are very different things. Apparently, teachers are unable to accomplish what they would like to be able to do.

Why don't mathematics teachers reach their own goals for teaching higher order strategies? If we knew how to evaluate higher order mathematical thinking, this question could be answered. Successful teaching of anything, including higher order thinking strategies in mathematics, is dependent upon the ability to determine the degree to which it has been learned. Valid and usable tests that speak to this concern can provide an impetus for teaching higher order thinking strategies.

Classroom Observation Instrument

In order to determine what is essential for promoting the learning of higher-order thinking skills, we selected the three primary classroom processes as defined by Peterson (1988):

(1) an emphasis on meaning and understanding (MAU);

(2) encouragement of student autonomy and persistence (SAT); and

(3) direct teaching of higher order cognitive strategies (HOS)

as the subheadings. The classroom observation instrument was developed to focus on classroom interactions that reflect these three processes. The observation form is shown in Table 1.

The first twelve items dealt with the emphasis on meaning and understanding as illustrated by classroom structure and curriculum implementation. The second group of seven items dealt with the student attitudes and interests in mathematics as reflected through the encouragement of student autonomy and persistence. The third group of seven items provide an indication of whether direct teaching of higher order cognitive strategies is implemented and whether the innovative testing approaches have an impact on the way students view mathematics. The indication as to how often the students work independently, solve problems in groups, use calculators or computers, or do hands-on activities cannot be directly addressed in the observations since some lessons lend themselves to these activities and some do not. In other words, the observer did not expect to have the opportunity to rate all items in each observation. However, over the three observations we were able to rate all of the items.

Table 1

Innovative Mathematics Assessment Project Observation Form

Teacher: ____________________ Date: ________

Class: ____________________ Time: ________

Emphasis on meaning and understanding:	Rating	Comments
Communicates that math problems cannot always be solved quickly		
Communicates that some problems have more than one answer		
Focuses on what students do know rather than what they don't know		
Uses informal assessment to provide feedback to students		
Mathematics is useful and makes sense		
Mathematical processes are used in context rather than in isolation		
Emphasizes understanding of mathematical concepts		
Provides opportunities to restate and formulate problems		
Provides opportunities to ask questions, consider different possibilities		
Mathematics is expressed through pictures, diagrams, graphs, words, symbols, or numerical examples		
Uses a variety of mathematical tools, models, manipulatives, calculators, or computers		
Provides opportunities for students to plan, invent, or design mathematical ideas, projects, activities, or products		

Encouragement of students autonomy and persistence:	Rating	Comments
Students learn at their own pace		
Students who perform with difficulty are not labeled as failures		
All students are expected to be able to learn mathematics		
Students work on extended assignments or investigations		
Speed is not an important factor in determining students' achievement		
Students are encouraged to think, be persistent, and self-directed		
Students work together to develop mathematical thinking skills		

Direct teaching of higher order cognitive strategies:	Rating	Comments
Teacher helps students to formulate and refine hypotheses		
Opportunities are given for collecting and organizing data and information		
Teacher helps students to learn and practice a variety of strategies for doing mathematics		
Teacher encourages students to reflect their own problem solving methods and strategies		
Students are asked to explain concepts orally or in writing		
Opportunities are given to work with open ended or poorly defined real-life problems		
Students are provided situations in which they enjoy doing mathematics		

In determining the ratings the observer and the teacher rated each item independently using a six level process based scale that addressed the items in approximately the following terminology (each teacher had the option to revise the meaning/rating if it did not serve to enhance understanding of the process):

0	Not present
1	Implied but not overtly present
2	Present but not developed
3	Present and used
4	Used in an insightful manner; room to expand
5	Developed and used; understanding obvious

The more subtle classroom activities such as "the refinement of hypotheses" or "enjoyment of doing mathematics" were more difficult to recognize since they can not always be recognized within one limited classroom visit. The interview that followed each observation gave the teacher the pportunity to explain how the lesson had (or would) develop over a period of days. Likewise, over a series of visits, the environment of the classroom became stable and these items began to be more obvious.

METHODOLOGY

The IMAP project observations began in October 1991 and ended in May 1992. The first step was to make an initial classroom visit and observation; meet with the Principal; and pick up the first student questionnaire. The observation and student survey were then repeated two more times.

Eighteen (18) teachers were observed during the school year. Before the first observation it was decided to video tape the classroom lesson for reference at a later time. By videotaping the lesson we could "systematically" check to see if the classroom experience gave the same impression via video tape as the actual observation did. The video tape also served as a checkpoint for correlating the comments the teacher put on their observation form. This also gave us a way to compare the three stages of observation for each teacher. An initial set of observations was done prior to implementation efforts to establish baseline data on teaching strategies. The observation schedule was arranged at the convenience of the teachers within the first months of the school year. The first observation (see Table 2) was scheduled with a block of time for the observer to visit with the Principal. The objective of this visit was to establish a working rapport within the school as well as determine the level of support the Principal would give the project. In most cases, the principal was extremely supportive and interested in the project.

TABLE 2

OBSERVATION 1: Schedule and Class Topics

Date	Observation	Teacher	Topic
10/31 1	Houston	BD	Isosceles Triangle
		FH	Polynomials
		OA	Sign Rules (Computers)
11/01 1	Rockdale	EB	Patterns
11/04 1	Bryan	SS	Test Review (Angles I/E)
11/05 1	College Station	CO	Measure of angles of triangle
11/06 1	Moody	JA	One variable equations
11/12 1	College Station	RB	Mean/Median
		GH	Multiplication 2 by 1
11/13 1	Calvert	DG	Geometry: point, line, etc.
11/14 1	Houston	VC	Multiplication
	Sugarland	GB	Odds/evens
		DS	Odds/evens
11/19 1	Montgomery	KN	Tessellation
11/21 1	La Grange	DA	Slope intercept
		CZ	Slope intercept
11/22 1	Austin	BR	Stock market graphs
11/26 1	Huntsville	CS	Polygons/definitions

The second observation (Table 3) was the first lesson that was observed in which the teachers were experimenting with the treatment plan they had proposed for alternative assessment. After each observation a block of time was reserved so that the observer and teacher could discuss their respective ratings. This became a time in which the teachers often expressed their joys as well as their frustrations with the treatment they were implementing.

TABLE 3

OBSERVATION 2: Schedule and Class Topics

Date	Observation	Teacher	Topic
02/13 2	Houston	BD	Area (Geoboards)
		FH	System of equations
		OA	Slope/graphing
02/06 2	College Station	RB	Area/ Perimeter
		GH	Writing word problems
02/07 2	La Grange	DA	Absolute values
		CZ	Absolute values/graphs
02/18 2	Bryan	SS	Special right triangles
2	College Station	CO	Statistics; ratio
02/20 2	Rockdale	EB	Place value
02/27 2	Huntsville	CS	Measurements/fractions
03/03 2	Montgomery	KN	Bisectors
03/05 2	Austin	BR	Review
03/10 2	Moody	JA	Slope
03/24 2	Houston	VC	Word problem skills
	Sugarland	GB	Angles; plane figures
		DS	Volume
03/26 2	Calvert	DG	Fractions

The third observation (Table 4) came at the end of the school year for the students as well as the teachers. By this time, they had completed the bimonthly class meetings on assessment and had tried a number of approaches to their respective plans of alternative assessment. The block of time after each observation was used to interview the participant in regards to their accomplishments or lack of accomplishments during the project. The teachers were extremely reflective.

TABLE 4

OBSERVATION 3: Schedule and Class Topics

Date	Observation	Teacher	Topic
05/01 3	Rockdale	EB	Angle = 180; 360
05/05 3	Montgomery	KN	Area
05/07 3	Houston	BD	Slope
		FH	Radicals
		OA	Factoring
05/11 3	Austin	BR	Stem & Leaf plots
05/12 3	Bryan	CO	Graphing of equations
05/13 3	Houston	VC	Capacity
	Sugarland	GB	geometry
		DS	school conflict
05/14 3	Huntsville	CS	Problem solving skills
05/15 3	La Grange	DA	Out ill
		CZ	Graphing inequalities
05/18 3	Moody	JA	Slope and y-intercept
05/20 3	College Station	RB	Bridge Building Project
		GH	recording bad
05/21 3	Calvert	DG	Multiplication two digit #s
05/22 3	Bryan	SS	Polyhedral

Analysis of data

Since we were using the same observation form throughout the project, we recruited an objective observer (mathematics major) to view and evaluate randomly selected video tapes. The independent observer selected fourteen (Table 5) of the observation tapes and over a period of three weeks viewed and scored the lessons. The scores of the independent observer and the classroom observer were then compared to determine an estimate of the reliability of the ratings.

Table 5

Independent Rater Selections

Observation	Independent Rater Selections Number of classes
1	5
2	4
3	5
Total	14

On each item the raters were judged to agree if their ratings were within one point on the 6 point scale. The degree to which the observation scale yields agreement by the two raters was at 88% for *Meaning and Understanding*; 92% for *Autonomy and persistence*; and 87% for *Higher Order Strategies* This gave an overall 98% agreement for the total of the three areas. Therefore, it is reasonable to assume that each repetition of the observation instrument to the same situation will yield similar results.

At the end of each class observation, the teacher being observed reflected upon the lesson by completing the same instrument as a self-evaluation measure. The percentage of agreement between the teacher and the observer ranged from 58% to 100% in *Meaning and Understanding*; 29% to 100% in *Autonomy and persistence* and *Higher Order Thinking* ; and 42% to 100% overall. If we dismiss the outlier in each category the percentage is 66, 57, 62 respectively to 100%.

The classroom observation data was analyzed to determine whether the work on assessment produced: (1) effects on the assessment strategies teachers use in the classroom, and (2) increased use of teaching approaches that enhance higher-order mathematical learning.

A one-way repeated measures analysis of variance of the observation data (Table 6) provided trend information on the nature of change in teaching processes brought about by the assessment training. Alpha levels of .05 and .01 were used to determine significance of F ratios.

TABLE 6

Summary of Analysis of Variance

SOURCE	DF	Meaning and Understanding		Student Autonomy and Persistence		Higher Order Thinking Skills		Total Overall	
		F	Pr	F	PR	F	PR	F	PR
Time of observation	2	5.67	0.0067	3.23	0.0499	5.02	0.0112	6.08	0.0049
Grade level	2	3.64	0.0351	3.09	0.0563	7.16	0.0022	[illegible].07	0.0107
Time*Grade	4	0.05	0.9951	0.04	0.9971	0.27	0.8935	0.09	0.9861

Since the observation and the grade level are studied independently the results were taken from the Type III summary.

The MAU F value for observations 1, 2, and 3 is 5.67 with a probability below alpha, indicating that there was significant change from observation 1 to 2 to 3. Likewise, the F value for grade levels (3.64) was below .05 indicating that there was a significant difference in MAU from each grade level (elementary to middle school to high school). There was no indicated interaction of the two variables.

The SAT F value for observations 1, 2, and 3 is 3.23 with a probability of 0.0317, indicating that there was significant change from observation 1 to 2 to 3. Likewise, the F value for grade levels (3.09) almost reached the alpha level indicating that there was a near significant difference in SAT from each grade level (elementary to middle school to high school). There was no indicated interaction of the two variables.

The HOS F value for observations 1, 2, and 3 is 5.02 with a probability below alpha indicating that there was significant change from observation 1 to 2 to 3. Likewise, the F value for grade levels (7.16) was well below alpha indicating that there was a highly significant difference in the direct teaching of higher order cognitive strategies between grade levels (elementary to middle school to high school). There was no indicated interaction of the two variables, time and grade.

By graphing the means of each observation within grade level for *Meaning and Understanding* (Figure 1) the elementary and high school means follow the same pattern even if there is a difference in the level of meaning and understanding that is emphasized.

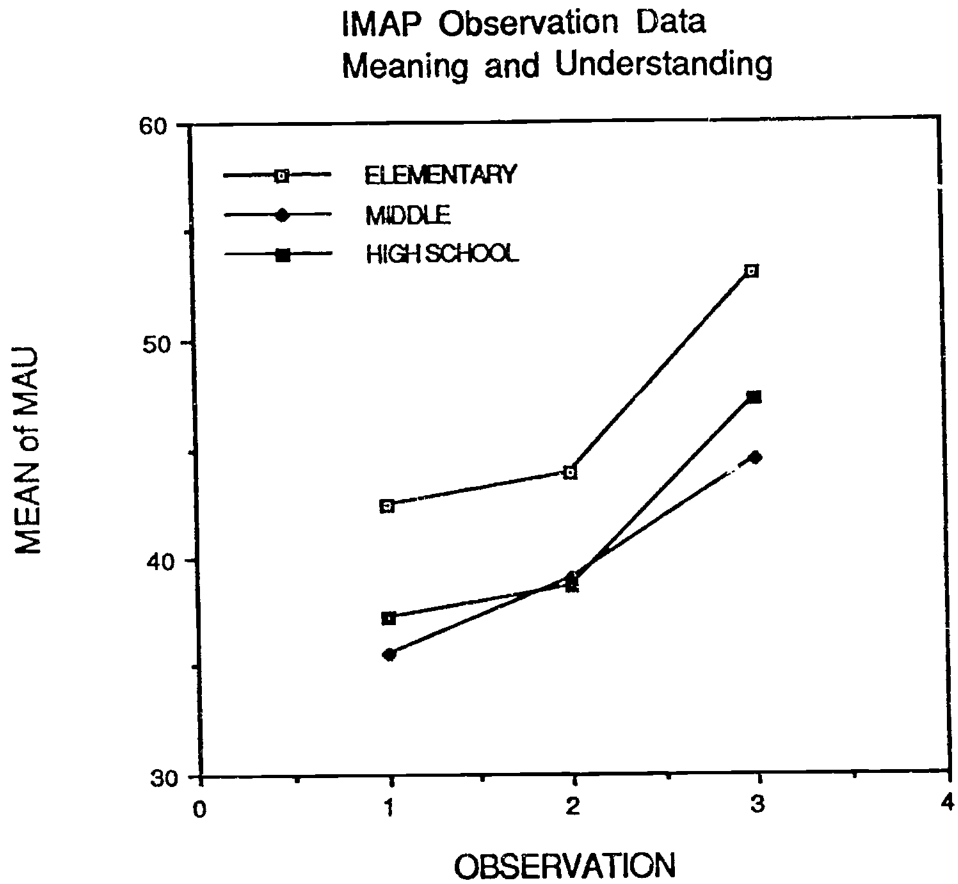

Figure 1.

Graph of cell means for MAU at three observations.

Meaning and Understanding was present and used in all of the classes, and over the period of the project the teachers increased their expertise at varying levels with the elementary and high school teachers showing a greater rate of increase, where the middle school had a steady linear increase in MAU.

By graphing the means of each observation within grade level for *Encouragement of student autonomy and persistence* (Figure 2) the elementary, middle school, and high school means follow the same pattern even if there is a difference in the level of student autonomy and persistence that is emphasized. It is also clear that the level of student autonomy and persistence increased fairly steadily over the period of the treatment regardless of the grade level.

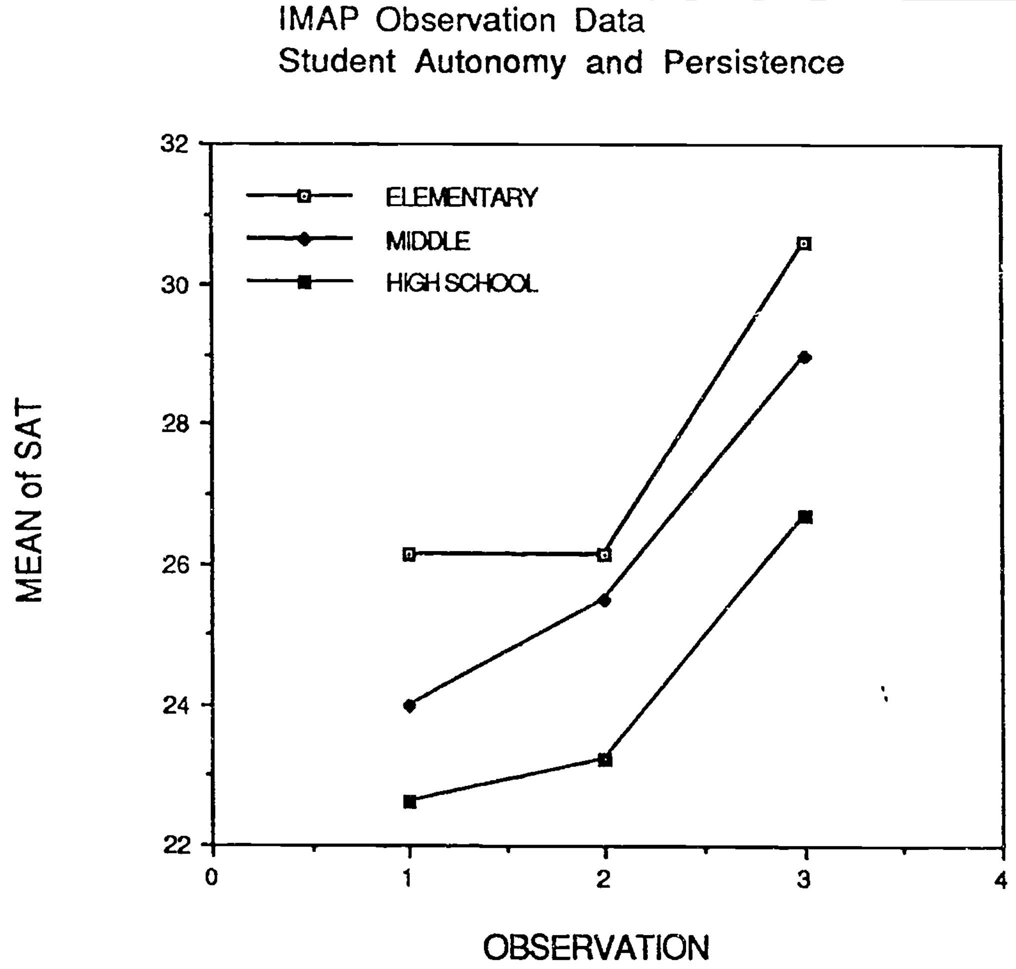

Figure 2.

Graph of cell means for SAT of three observations.

The *Encouragement of student autonomy and persistence,* although at a higher level in the elementary, this area tended to hold steady as the treatment began, where as the middle and high school indicated a steady increase of student autonomy and persistence. This may be due to the nature of the high school environment and the student attitude more than the lack of innovative measures.

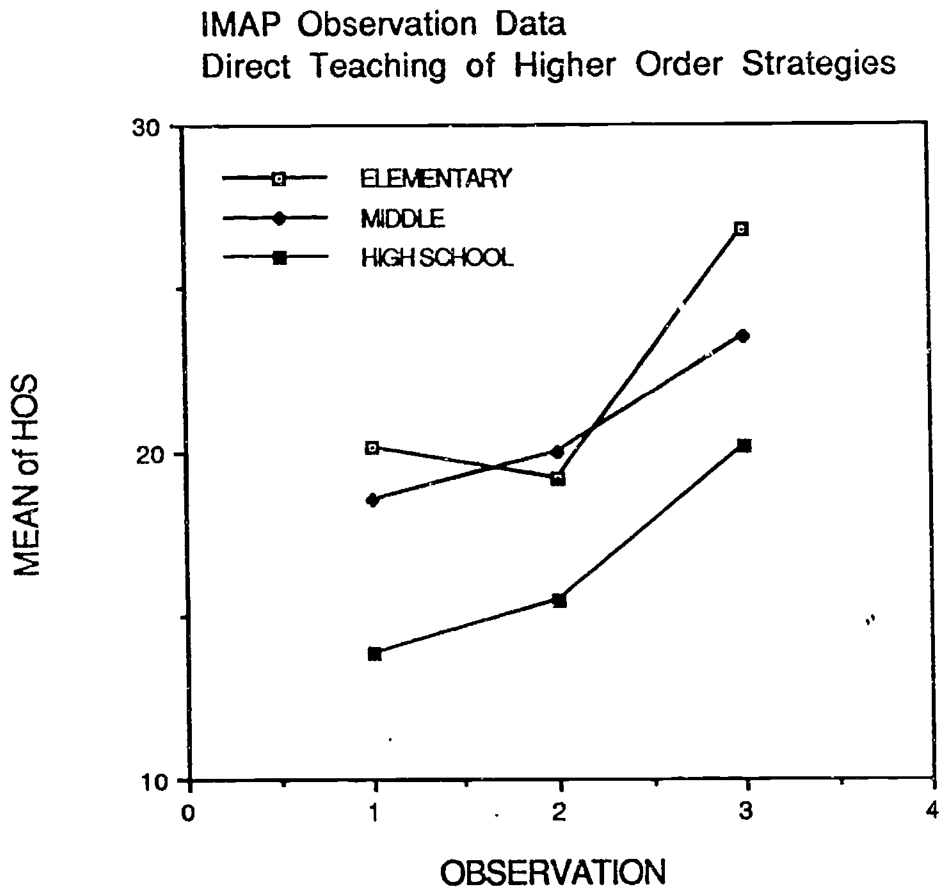

Figure 3.

Graph of cell means for HOS of three observations.

By graphing the means of each observation within grade level for the *Direct teaching of higher order cognitive strategies* (Figure 3) the middle school and high school means follow the same pattern even if there is a difference in the level of HOS that is emphasized. It is also clear that the level of HOS decreased in the elementary level before taking a drastic upward swing at the end of the treatment. The middle school paralleled the high school over the period of the treatment even though the middle school had a higher level of emphasis on HOS.

The *Direct teaching of higher order cognitive strategies* showed substantial increase at all levels by the third observation. This may be due to the nature of mathematics in that the skills needed have been mastered by the end of the school year and the teachers are involved in more project and problem solving activities.

DISCUSSION

In order to see if one area contributed more than another area each item was considered individually over the observation cycle. The high school teachers seemed to assume *that math problems could not always be solved quickly* and began to communicate it to the class in a verbal manner with statements such as "take your time"; "you may not see it the first time"; and "think about what you are doing". Only a couple of teachers actually gave problems that were identified as on-going or multi-day tasks. Likewise, the farther into the year of mathematics, the less indication we see that *some problems have more than one answer.* The more complicated the algebra/geometry became, the less flexibility was seen. All of the high school teachers were at a high level of *focusing on what students do know* and they continued to improve on this item. There was a definite increase in *informal assessment to provide feedback to students* however the attention to this item was more intense in the middle of the semester than at the end. This may be a direct result of our method of grading and evaluation for continuing (or not) in a subject area. There was no overall continuity in *mathematics is useful and makes sense.* Some of the teachers actually had significant decreases in this area. In the interview process the most commonly heard comment was "I just do not know what this is needed (good) for. My background in mathematics history is very weak. I spend all of my time in college learning how to make the formulas work." A number of the teachers were making an effort to increase their own background along with the students by using sponge activities that had to do with history and/or the use of particular areas of mathematics. The majority of the high school teachers were conscious of *using the processes in context rather than isolation* even if it was only a brief introduction of the environment. The high school teachers were already *providing opportunities to restate and formulate problems; to ask questions, consider different possibilities* and they continued to improve in these areas. The largest increase in activity came in the areas of *mathematics expressed through pictures, diagrams, graphs, words, symbols, and numerical examples; use of a variety of tools, models, manipulatives, calculators, or computers* for the enhancement of meaning and understanding. The high school teachers did not give as many *opportunities to plan, invent, or design mathematical ideals, projects, activities, or products* as the elementary and middle school teachers did. One teacher explained "...the classes are just too large; and the students think that most of the projects are elementary. It just isn't worth getting behind in the text to provide the opportunity."

In the area of *encouragement of student autonomy and persistence* three items were negativity influenced. *Speed* was constantly an important factor if for no other reason - the length of the class period created boundaries at the high school level. Even though the teachers were above average in *encouraging the students to think, be persistent, and self-directed* this item consistently dropped across the teachers' observations but was always present. The lack of a clear pattern in *students working together to develop mathematical thinking skills* at the high school level may be due to the lack of working in groups and/or the desire to guard ones' work when given an opportunity to work together. One of the algebra teachers commented, "I give them opportunities to work together, but most of them feel that they are cheating if someone else (other than the teacher) helps them."

The *direct teaching of higher order cognitive strategies* was the easiest to recognize and record. The high school teachers did more of direct teaching as we approached the end of the year. The only areas not showing a consistent increase was *the encouragement of the student to reflect their own problem solving methods and strategies* and *the students being asked to explain concepts orally or in writing.* The only explanation given in the interviews was "we just don't have any more time and there are too many in the class to hear from them all".

SUMMARY

The main components of assessment - situation, response, analysis, and interpretation - provide information to help the teacher work with the students in the hope that both will gain a greater understanding of the far-reaching implications of mathematics. The purpose for the assessment must be in alignment. From the results of this project many questions have come to the forefront such as:

Does the assessment of mathematics differ from the assessment of other content areas to the extent that a distinct and separate theory is meaningful?

Can the techniques of assessment in mathematics be implemented (integrated) into other areas so that the lines of demarcation can be erased along with the "fear". If we accept Webb's definition that "The term mathematical assessment refers to the comprehensive accounting of an individual's or group's functioning within mathematics or in the application of mathematics", then our concept of assessment must be considered from a variety of perspectives. The teacher will have to be able to provide feedback to the student not in the terms of a grade but in terms of knowledge gained and paths to take. Which means: How do we develop tests that are a measure of meaning and understanding rather than just skills? How do we select the best form of analysis for a specific assessment situation? How do researchers interpret and understand the interaction among the various elements? How do we embed assessment in order to make use of limited class time and not get caught by the "grade axe"? The practical procedures for doing this are yet to be developed. More research needs to be done to investigate the actual practice and impact of a variety of assessment techniques in the classroom. Tests do influence teachers and what they teach. Little is known about aggregating mathematics assessment data and what form of analyses are needed to derive maximum information from one or more assessments. What is the number of assessment situations needed to understand what a student knows? Why do we change the methods of imparting information when we move from the elementary to middle school to high school? Do the way the students learn change? Can assessment changes point the way toward effective and efficient teacher in-service training? Rather than doing separate in-service workshops on instructional approaches, and on innovative assessment, we can develop approaches that build on each other.

Mathematics teachers in the future will seldom tell students to put away their books and notes to take a test. Assessment will be an ongoing and integrated part of teaching, so that students can use everything to their advantage. Students will have the opportunity to provide input to their assessment records, making sure that their best work, produced without pressure and anxiety, is included in the evaluation process.

References

Charles, R. I. & Silver, E. A. (1989). *The teaching and assessing of mathematical problem solving.* Reston, VA: National Council of Teachers of Mathematics.

Council ot Chief State School Officers (1988). *Assessing mathematics in 1990 by the National Assessment of Educational Progress.* Washington, DC: Author.

Crosswhite, F. J., Dossey, J. A., Swafford, J. O., McKnight, C. C. & Cooney, T. J. (1986). *Second international mathematics study, Summary report for the United States.* Champaign, IL: University of Illinois.

Dossey, J. A., Mullis, I. V., Lindquist, M. M., & Chambers, D. L. (1988). *The mathematics report card.* Princeton, NJ: Educational Testing Service.

Foxman, D. & Mitchell, P. (1983). Assessing mathematics: APU framework and modes of assessment. *Mathematics in schools*, 12(4), 1-5.

Kulm, G. (1990). *Assessing higher order thinking in mathematics.* Washington, DC: American Association for the Advancement of Science.

Matthews, W., Carpenter T. P., Lindquist, M. M., & Silver, E.A. (1984) The Third National Assessment: Minorities and mathematics. *Mathematics Teacher*, 74(7) 165-171.

National Assessment of Educational Progress. (1987). *Learning by doing. A manual for teaching and assessing higher-order thinking in science and mathematics.* Princeton, NJ: National Assessment of Educational Progress.

National Council of Teachers of Mathematics. (1992). *Curriculum and evaluation standards for school mathematics* Addenda Series, Grades 9-12 A Core Curriculum Making Mathematics Count for Everyone. Reston, VA: NCTM.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics.* Reston, VA: NCTM.

Oakes, J. (1987). Race, class, and school responses to `ability': Interactive influences on math and science outcomes. Paper presented at annual meeting of The American Educational Research Association, Washington, DC.

Pandey, T. (1990). Power items and the alignment of curriculum and assessment. In G. Kulm, (Ed.), *Assessing higher order thinking in mathematics.* Washington, DC: American Association for the Advancement of Science.

Peterson, P. L. (1988). Teaching for higher-order thinking in mathematics: The challenge for the next decade. In D. A. Grouws, T. J. Cooney & D. Jones (Eds.) *Effective mathematics teaching.* Washington, DC: National Council of Teachers of Mathematics.

Romberg, T., Zarinnia, A., & Collis, K. (1990). A new world view of mathematics assessment. In G. Kulm, (Ed.), *Assessing higher order thinking in mathematics.* Washington, DC: American

Association for the Advancement of Science.

Stenmark, J. K. (1991). *Mathematics assessment: Myths, models, good questions, and practical suggestions.* Reston, VA: NCTM.

Stenmark, J. K. (1989). *Assessment alternatives in mathematics* Berkeley, CA: University of California.

Strong, D. (1990). Calculators and mathematics assessment. In G. Kulm, (Ed.), *Assessing higher order thinking in mathematics.* Washington, DC: American Association for the Advancement of Science.

Webb, N. L. (1975). An exploration of mathematical problem-solving processes (Doctoral dissertation, Stanford University). *Dissertation Abstract International, 36,* 2689A.